Reprinted from JOURNAL OF METEOROLOGY, Vol. 9, No. 3, June, 1952, pp. 172–175
Printed in U. S. A.

STATISTICAL PROPERTIES OF PRESSURE CHANGE ALOFT

By William Kaciak and Patricia A. Langwell

New York University[1]

(Original manuscript received 5 January 1952; revised manuscript received 20 February 1952)

ABSTRACT

The basis of many statistical tests is the assumption that the data are drawn from a normal population. The change in pressure over one to ten days at 3000 dynamic meters, at 50°N and 70°W, is tested, and the circumstances under which such pressure changes form a normal distribution are described. Correlograms of the pressure change at this location are calculated, by season, for the period 1934 through 1940. From this, a population of near-zero autocorrelation is defined. For this limited population, it is demonstrated that the mean and skewness are zero, and that the kurtosis is three. Next, the distribution is shown to be normal by the χ^2 test. Finally, an estimate is given of the size of sample necessary to eliminate the synoptic effect.

1. Introduction

In connection with a study of the variation of pressure with geomagnetic activity, the statistical properties of the change in pressure at 50°N and 70°W have been tested. The pressures were read from the *Historical weather maps*. The choice of location was dictated by the following considerations. The relationship between pressure and geomagnetic activity improves with increasing latitude. Therefore, 50°N was chosen as the northernmost latitude for which sufficient data are available. The longitude 70°W is nearest to a satisfactory network of stations, over North America, without being too close to orographic disturbances. Autocorrelations have been calculated for lags of one through ten days by season, and the circumstances under which the pressure change forms a normal distribution will be described.

2. Autocorrelation

Since the pressure patterns of any day resemble the pressure patterns of preceding and following days, it is reasonable to assume that a correlation exists between pressures on successive days. The bias due

[1] The research reported in this paper has been sponsored by the Geophysics Research Division of the Air Force Cambridge Research Center, under Contract AF 19(122)-49.

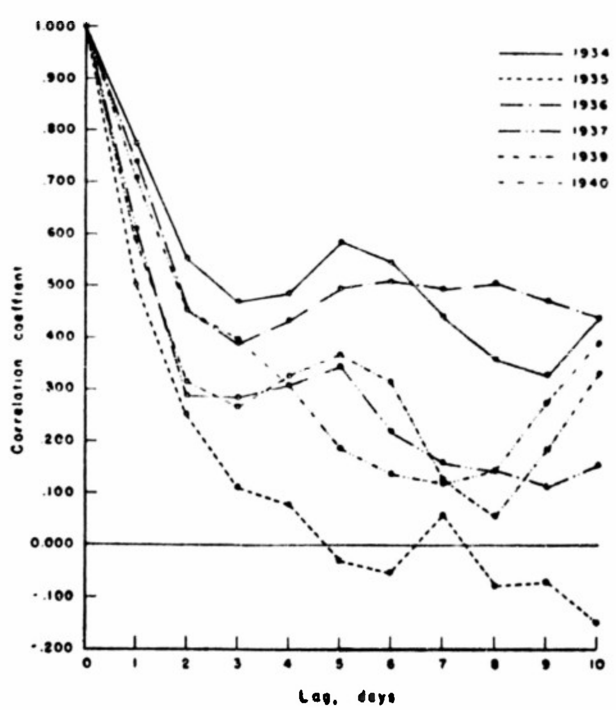

FIG. 1. Correlogram of 3000-gdm pressure in fall (September–November).

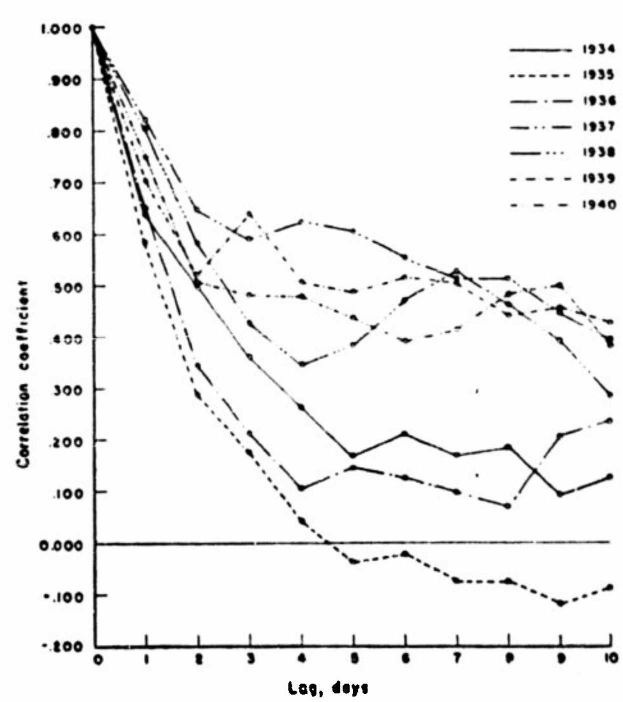

FIG. 2. Correlogram of 3000-gdm pressure in spring (March–May).

to this persistence is eliminated if the correlation between the members is zero. Namias (1947) computed this correlation for the point 40°N and 90°W at 10,000 ft, using a year's data during 1944–1945. The resultant curve is similar to the average of those below. Klein (1951) has combined a great deal of data and computed a spatial distribution of the correlation for a one-day lag.

Since the correlation may vary seasonally, the pressures at 50°N and 70°W for the period 1934 through 1940 are divided into four seasons; winter (December through February), spring (March through May), summer (June through August), and fall (September through November). Correlograms are constructed for each season, for lags of one through ten days. The correlation coefficient is given by

$$r_e = \frac{\sum_{i=1}^{n-e} P_i P_{i+e} - \frac{1}{n-e}\sum_{i=1}^{n-e} P_i \sum_{i=1}^{n-e} P_{i+e}}{\left\{\left[\sum_{i=1}^{n-e} P_i^2 - \frac{1}{n-e}\left(\sum_{i=1}^{n-e} P_i\right)^2\right]\left[\sum_{i=1}^{n-e} P_{i+e}^2 - \frac{1}{n-e}\left(\sum_{i=1}^{n-e} P_{i+e}\right)^2\right]\right\}^{\frac{1}{2}}}, \qquad (1)$$

where P_i is the pressure on any given day, P_{i+e} the pressure e days later, and n the number of observations. The results for the four seasons are shown in figs. 1, 2, 3, and 4. In the fall and spring, the correlations do not reduce to zero. They drop off rapidly for lags one and two, and then level off at an average value of 0.25. There is a great deal of scatter from year to year, with the level value varying from 0.50 to −0.10. The failure of the correlation coefficient to reduce to zero is probably a result of the seasonal trend of rising pressures in the spring and falling pressures in the fall.

In the winter, the curves are remarkably similar for lags one through five. All five yearly correlations reduce to zero between four and five. For lags five through ten there is considerable variation from year to year, with the average near zero. In the summer the shape of the curves is similar to that of the curves of the spring or fall, except that the correlation levels off at a value near zero. As a result of these calculations, it may be stated that the seasonal effect and the persistence effects are probably eliminated if changes of pressure over three days or more, in the summer or winter, are considered.

3. Normality of the frequency distribution

The population is now limited to the seasons and lags described at the end of the last section. The pressure on the eth day minus the pressure on the zero day is defined as ΔP_e, and this quantity is computed for e = 3, 4, 5, and 6. Class intervals of 5

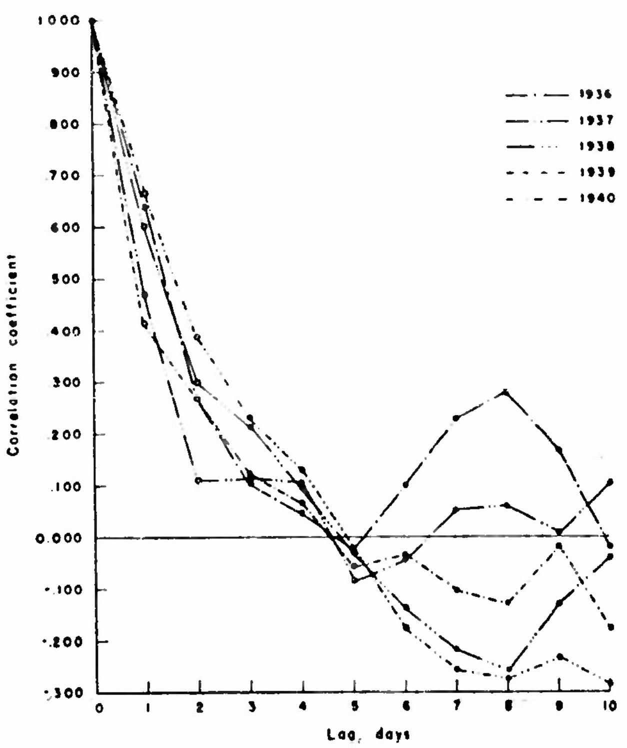

FIG. 3. Correlogram of 3000-gdm pressure in winter (December–February).

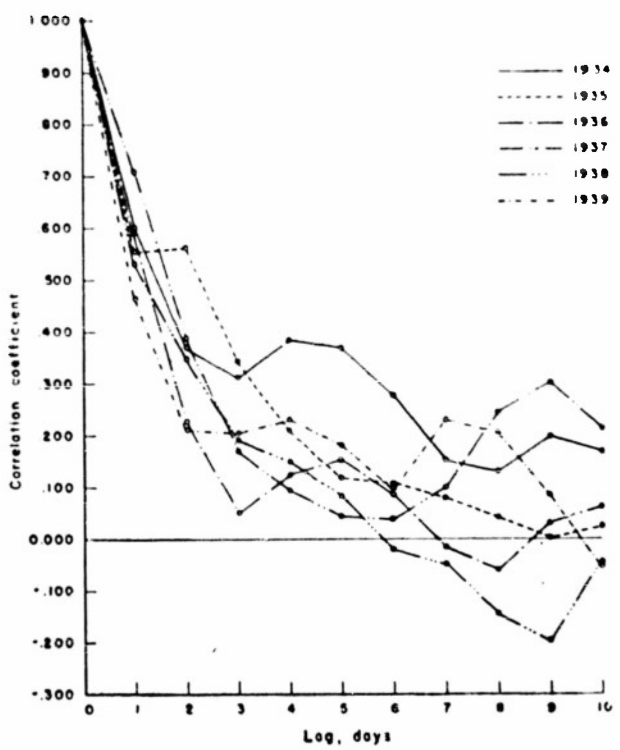

FIG. 4. Correlogram of 3000-gdm pressure in summer (June–August).

mb are chosen, with the zero interval containing pressure changes of -2.5 to $+2.5$ mb. Such a division into class intervals facilitates calculations for large samples, and is especially helpful in a later application of the χ^2 test. Frequency distributions of the ΔP_t are tabulated, and the mean, and the second, third, and fourth moments about the mean are calculated (Hoel, 1947).

The mean is

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{h} x_i f_i, \tag{2}$$

where x_i is the midpoint of the class interval, f_i the frequency of occurrence in the class interval, h the number of intervals, and $n = \sum_{i=1}^{h} f_i$ is the total number of cases.

The kth moment is

$$m_k = \frac{1}{n} \sum_{i=1}^{h} (x_i - \bar{x})^k f_i. \tag{3}$$

The 2nd moment, m_2, is the variance. Its root, $m_2^{\frac{1}{2}}$, is the standard deviation.

The third moment, m_3, is a measure of the skewness. However, to obtain a pure number independent of the units used, the third moment is divided by the cube of the standard deviation. If this measure of skewness is denoted as a_3,

$$a_3 = m_3/m_2^{\frac{3}{2}}. \tag{4}$$

If the distribution is normal, a_3 is zero, since a normal curve is symmetrical about the mean.

The fourth moment, m_4, is a measure of kurtosis. To obtain a dimensionless quantity, a_4, m_4 is divided by the square of m_2. Thus,

$$a_4 = m_4/m_2^2. \tag{5}$$

In a normal distribution, $a_4 = 3$.

The mean and the moments about the mean are tabulated in table 1. From table 1, it can be observed that the actual values of the mean and the moments of ΔP_t are in good agreement with the theoretical values of a normal distribution. However, in the winter $\bar{x}$ has a slight tendency to be negative, while in the summer $\bar{x}$ has a slight tendency to be positive. This would indicate that, on the average, the pressure drops slightly in winter and rises slightly in summer as time progresses. The standard deviation $m_2^{\frac{1}{2}}$ is larger in winter than in summer. This is due to the larger range of pressure fluctuation in winter. The measures of skewness and kurtosis are so close to the theoretical values that no comment seems necessary.

TABLE 1. The mean and the moments about the mean for ΔP_t in the winter and summer.

	$\bar{x}$ (mb)	$m_2^{\frac{1}{2}}$ (mb)	a_3	a_4
		Winter		
ΔP_2	-0.2	11.5	-0.1	2.9
ΔP_4	-0.2	12.0	0.0	2.8
ΔP_6	-0.4	12.0	0.0	2.8
ΔP_8	-0.3	12.8	0.0	2.9
		Summer		
ΔP_2	0.3	7.4	0.0	3.1
ΔP_4	0.3	7.8	0.0	3.0
ΔP_6	0.4	7.6	0.0	3.1
ΔP_8	0.6	7.8	0.1	2.9

Since all the tests of central tendency are satisfactory, an additional test of the distribution away from the mean is applied. The χ^2 test determines whether there is a significant difference between the actual and a theoretical distribution.

$$\chi^2 = \sum_{i=1}^{h} \frac{[f_a(x) - f_T(x)]^2}{f_T(x)}, \tag{6}$$

where $f_a(x)$ is the actual frequency, and $f_T(x)$ is the theoretical frequency. Since, in our case, the theoretical frequency is the normal distribution,

$$f_T(x) = \frac{1}{\sigma(2\pi)^{\frac{1}{2}}} \exp\left[-\tfrac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2\right], \tag{7}$$

where x is the value of the variate, $\bar{x}$ the mean of the variate, and σ the standard deviation. It is assumed that the mean and standard deviation of the sample are representative of the population as a whole. Since these two quantities have been calculated above, it is possible to tabulate the theoretical distribution without difficulty.

It can be seen from (6) that a zero value for χ^2 corresponds to an exact agreement between the actual and theoretical values, whereas increasing values of χ^2 correspond to increasingly poor agreement. Therefore, a value of χ^2 is selected, χ_0^2, as a critical value for judging significance, such that the probability of the true distribution being normal even though $\chi^2 > \chi_0^2$ is 0.05. The χ^2 distribution function depends only on the parameter ν, the number of degrees of freedom. The degrees of freedom in this case equal the number of class intervals minus three. By knowing the degrees of freedom, it is possible to compare calculated values of χ^2 with χ_0^2 for the 5-per cent level (Hoel, 1947, table III, p. 246). If $\chi^2 > \chi_0^2$, the actual observations are significantly different from the theoretical, and the ΔP_t frequencies are not normally distributed.

The results of the χ^2 test for the goodness of fit of the actual ΔP_t distributions are shown in table 2. In every case, in both winter and summer, $\chi^2 < \chi_0^2$. However, inasmuch as the χ^2 test is relatively insensitive, it is necessary to consider how well within the limit the computed values of χ^2 lie. It can be seen that the winter values of χ^2 are proportionately less. As a result, even though the summer ΔP_t meet the requirement of the χ^2 test, they must be used with caution. Moreover, the smallest sample used in this test contains 450 items. This is larger than is often available to the statistical meteorologist. To discover whether a small sample meets the requirement of approaching the normal distribution, a sample of 25 days was selected with the aid of a table of random numbers (Snedecor, 1946). The 25 days were selected from the 720 available during the period 1932–1940. The pressure on the key day is subtracted from the pressure four days later, and the average ΔP_4 of one sample is calculated.

The hypothesis is made that $\overline{\Delta P_4}$ is zero, within the error of the sample. The 5-per cent confidence level is again chosen. The standard deviation of the means of samples of size n, σ_m, is calculated.

$$\sigma_m = \left[\frac{\sum x^2}{n^2} - \frac{\bar{x}^2}{n} \right]^{1/2}, \tag{8}$$

where the members of the sample take on the values x and have the mean $\bar{x}$. If the hypothesis is true, $|\bar{x}| < 2\sigma_m$. The results for two separate samples of 25 are:

		Winter	Summer
Sample (a):	$\overline{\Delta P_4}$	2.4	−2.7
	$2\sigma_m$	4.8	2.6
Sample (b):	$\overline{\Delta P_4}$	0.9	−1.9
	$2\sigma_m$	4.4	2.6

Since the hypothesis is not verified for one of the summer cases, the samples were enlarged to 49 by the choice of 24 additional cases. The results for the two samples of 49 are:

		Winter	Summer
Sample (a):	$\overline{\Delta P_4}$	0.6	−1.7
	$2\sigma_m$	3.2	2.2
Sample (b):	$\overline{\Delta P_4}$	−0.2	−1.3
	$2\sigma_m$	3.0	2.2

TABLE 2. χ^2 test for the goodness of fit of the ΔP_n distributions to the theoretical normal curve. $P[\chi^2 > \chi_0^2] = 0.05$.

	χ^2	χ_0^2	ν
	Winter		
ΔP_3	10.0	15.5	8
ΔP_4	9.4	16.9	9
ΔP_5	12.1	18.3	10
ΔP_6	6.5	18.3	10
	Summer		
ΔP_3	12.3	12.6	6
ΔP_4	10.4	12.6	6
ΔP_5	8.8	12.6	6
ΔP_6	9.3	12.6	6

For a sample of 49, the hypothesis is satisfied for both seasons.

4. Conclusions

For every test, the change in pressure over more than three days in the winter fulfills the requirements of a normal-distribution test. In the spring and fall, the autocorrelations of pressure do not reduce to zero, so that the moment and χ^2 test have not been applied. The change in pressure over more than three days in the summer satisfies the conditions for the normal distribution. However, to eliminate the effect of the weather situation, a larger sample is needed in the summer than in the winter.

REFERENCES

Hoel, P. G., 1947: *Introduction to mathematical statistics*, New York, John Wiley and Sons, 258 pp.

Klein, W. H., 1951: A hemispheric study of daily pressure variability at sea level and aloft. *J. Meteor.*, 8, 332–346.

Namias, J., 1947: *Extended forecasting by mean circulation methods*. Washington, U. S. Weather Bureau, 89 pp.

Snedecor, G. W., 1946: *Statistical methods*. Ames, Iowa State College Press, 485 pp.